Introduction to this Course: Multilevel Models for Longitudinal Data

- Topics:
  - What to expect this semester
  - Features of longitudinal data
  - Features of longitudinal models
  - What can MLM do for you?

What To Expect This Semester

- I believe that **everyone is capable** and **can significantly benefit**** from learning more types of quantitative methods (*especially* multilevel models)!

- **Philosophy:** Focus on accessibility + mastery learning
  - No anxiety-prone tasks (e.g., hand calculations, memorizing formulas)
  - No need to resort to AI for help! That's the job of me and the TAs!

- **Materials:** Unit = lecture + example(s); 7 units planned
  - **Lecture** slides present concepts—the (wordy) **what** and the **why**
  - **Example** documents: reinforce the concepts and demonstrate the **how using software**—STATA or R (SAS also still provided)
  - All available at the course website (hosted outside of ICON)

** ***Benefits*** *include but are not limited to: Better research, more authorship opportunities, and actual money*

What To Expect This Semester

- I will NOT:
  - Use infrequent high-stakes testing to assess your level of learning
  - Assume you know any models beyond the GLM to begin with!
    - But I will make connections to other modeling frameworks (MLM → SEM)

- I WILL:
  - Use **formative assessments (FA)** to help you figure out what you need to review (7 planned; 14 points for **completing them at all**)
  - Require **online homework (HW) assignments** that give you real-data practice (5 planned; 86 points for **completing them accurately**)
    - Complete demo "HW 0" for 2 points extra credit
    - All canned data this semester, but questions about your own analyses are welcome—and appreciated!—during class or office hours
  - Link research designs, data, questions, and models explicitly
    - If we don't cover the exact combination you need in class, just ask me—odds are good I have an example from elsewhere that I can send you!

More About the Course Requirements

- **Everything** is take-home, open-note, and untimed
- Late* work will be accepted (−2 for HW or −1 for FA, overall)
  - **Extensions granted if requested at least 2 weeks in advance*
  - HW due dates **may be pushed later** (to ensure approximately 1 week after covering the material before it's due), but never sooner
- **Formative assessments:** Big-picture questions to provide a structured review (will go over answers at the next class)
  - Also likely to include case studies and open-ended review prompts
  - Minimal feedback given, so please ask any remaining questions in class!
- **Homework assignments:** Practice doing data analysis
  - Based directly on examples given (no googling or ChatGPT required)
  - You will each have a unique dataset (made with a common story)
  - **Computation** sections: Instant feedback, infinite attempts
  - **Results** (interpretation) sections: Delayed feedback, single attempt (but repetition of concepts and vocabulary across the semester)

Our Other Responsibilities

- My job (besides providing materials and assignments):
  - **Answer questions** via email, in individual meetings, or in group-based zoom office hours—you can each work on homework during office hours and get (near) immediate assistance (and then keep working)

- Your job (in descending order of timely importance):
  - **Ask questions**—preferably in class, but any time is better than none
  - **Frequently review** the class material, focusing on mastering the vocabulary (words and notation), logic, and procedural skills
  - Don't wait until the last minute to start homework, and don't be afraid to **ask for help if you get stuck** on one thing for more than 15 minutes
    - Please email me a screenshot of your code+error so I can respond easily
  - **Do the readings** for a broader perspective and additional examples (best after lecture; readings are for the whole unit, not just that day)
  - **Practice** using the software to implement the techniques you are learning **on data you care about**—this will help you so much more!

More About Your Experience in this Class

- **Attendance**: Strongly recommended but not required
  - **You choose** (for any reason): In-person "roomer" or "zoomer"
  - **Masks** are welcome for in-person attendees
  - **Please do not attend in-person if you may be ill!**
  - You won't miss out: I will post **YouTube-hosted recordings** (audio + screenshare only) for each class at the course website
  - **Ask questions aloud or in the zoom chat window (+DM)** (even if you are attending class in person)

- **Changes** will be sent via email by 9 am on class days
  - I will change to zoom-only if I am sick!
  - I will change to zoom-only for dangerous weather
  - Nothing is more important than our health and safety…

Class-Sponsored Statistical Software

- To help address the needs of different Iowa degree programs, I will show examples using **STATA and R** software (and some **SAS**)
  - **STATA** (aka, Stata) = "Software for Statistics and Data Science"
  - **R** = free implementation of what was initially the "S" language
  - **SAS** = "Statistical Analysis System" (will appear in some examples)

- **Why not SPSS**? Because it doesn't have as much room to grow (and thus it isn't used in any other EMS advanced classes)
  - As in SPSS, drop-down windows can also generate syntax in STATA
  - SPSS could be used for some of our content (see my textbook website)

- **My story:** After SPSS, I became a heavy-duty **SAS enthusiast** who:
  - Picked up enough STATA initially to teach workshops using it, and I am learning it better now that I teach it in my classes
  - Is (begrudgingly) learning enough (base) R to add it to my classes, so it's possible you learned how to do the same things differently but correctly
  - So if you have **STATA or R tips**, please share them with me!
  - I also use Mplus, which will be used in a few bonus examples in this class

Which Program: STATA, R, or SAS?

- **I am assuming you know how to use at least one of these!**
  - Each is available (with VPN) in the free U Iowa Virtual Desktop
  - More programs = more "technical skills" for your CV; easier collaboration with colleagues (who may only know one program)
  - For intro handouts and videos, please see materials for PSQF 6243

- **To consider** when choosing which program to focus on:
  - Future use: R can be freely installed on your own machine; SAS has a free web-based SAS OnDemand; STATA install = $$$
  - **STATA** is popular in fields that use **large, weighted survey data** (e.g., sociology, political science, public health, EPLS at Iowa)
  - **R** will be used exclusively in classes by Drs. Aloe and Templin, and it has become increasingly mainstream, **but**:
    - R packages are only as good as their authors (so little quality control)
    - Syntax and capabilities are idiosyncratic to the packages (grrrrrr)
    - I have found incorrect or impossible results, even for "good" packages

This Semester: Longitudinal Multilevel Models

- **Background:** how this relates to what you already know
  - Concepts in longitudinal modeling (and why use MLMs)
  - Independent review of concepts from single-level linear models
  - Within-person analysis via Repeated Measures Analysis of Variance

- **"Unconditional" longitudinal models:** describing an outcome's pattern of means, variances, and covariances over time
  - Due to within-person fluctuation via alternative covariance structures
  - Due to within-person change via fixed and random effects of time (polynomial, piecewise, and exponential models)
  - This lengthy component can be skipped for other kinds of multilevel data

- **"Conditional" longitudinal models:** adding predictors
  - Time-invariant (level-2) cross-sectional predictors
  - Time-varying (level-1) predictors within multivariate MLMs ("M-SEM") will be addressed in the Spring 2025 PSQF 7375 advanced class instead

Data Requirements for Our Models

- A useful outcome variable:
  - Has an interval scale with plausibly continuous residuals*
    - A one-unit difference means the same thing across all scale points
    - **Other kinds of outcomes can be analyzed using generalized multilevel models instead, as addressed in the PSQF 6272 clustered MLM class*
  - Has scores* with the same meaning over observations
    - Includes meaning of construct and/or how items relate to the construct
    - Implies measurement invariance (and/or vertical equating) over time
    - **This class will address observed outcomes only; change over time in latent variables will be addressed in the Spring 2024 advanced class*
- FANCY MODELS STILL CANNOT SAVE BADLY MEASURED VARIABLES OR CONFOUNDED RESEARCH DESIGNS.

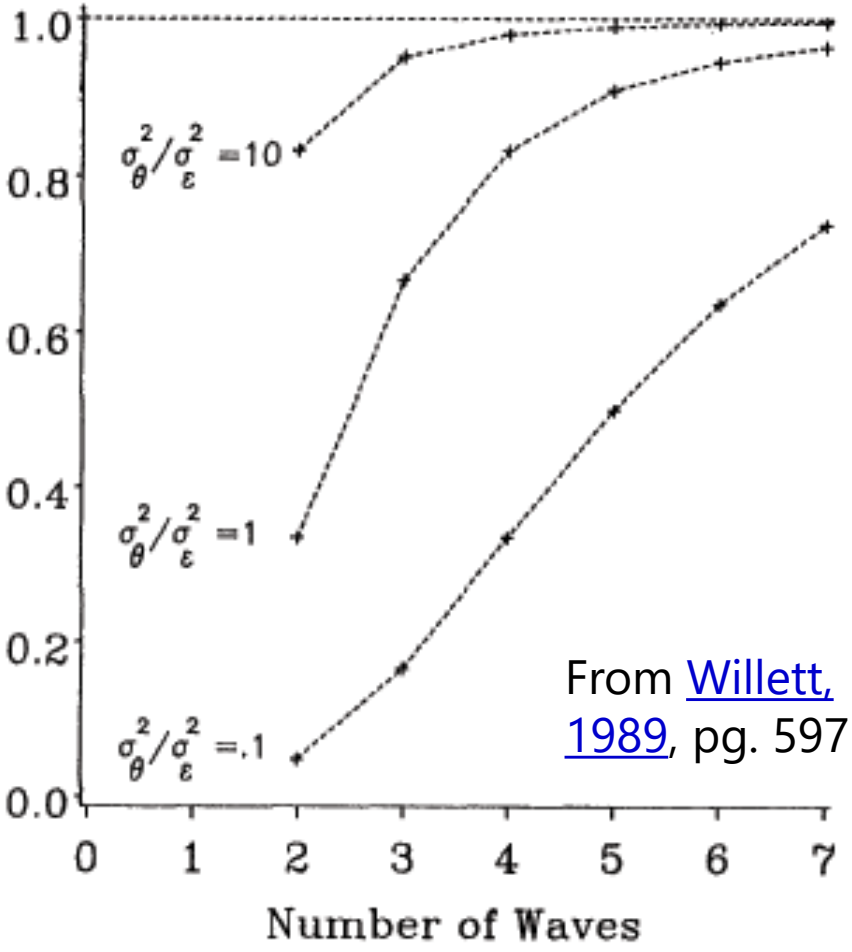
Requirements for Longitudinal Data

- Multiple repeated **observations** from same sampling "unit"
  - "Units" can be anything: persons, schools, countries, animals…
  - "Repeated" can span any length of time (milliseconds to days to years)
  - "Repeated" can also include trials: across items, conditions, etc.
- 2 occasions is the minimum, but just 2 can lead to problems:
  - Only 1 kind of change is observable (1 linear difference)
  - Can't distinguish "real" individual differences in change from error
  - Repeated measures ANOVA is just fine for 2 (complete) observations
    - Necessary assumption of "sphericity" is satisfied with only 2 observations
- More data is better (with diminishing returns)
  - More occasions → better description of the form of change; more power to show effects of time-varying predictors (measured repeatedly)
  - More units (persons) → better estimates of the amount of individual differences in change; better prediction of those individual differences
  - More items/stimuli/groups → more power to show effects of differences between items/stimuli/groups (i.e., for each dimension of sampling)

Power in Longitudinal Data

- **More occasions are better!**
  - Can examine more complex growth functions
  - Can get more reliable individual growth parameters
- **More units (people) are better!**
  - Can get more reliable estimates and prediction of individual differences

Reliability of Slopes (y-axis) by Signal-to-Noise Ratio and # Occasions



From Willett, 1989, pg. 597

Levels of Analysis in Longitudinal Data

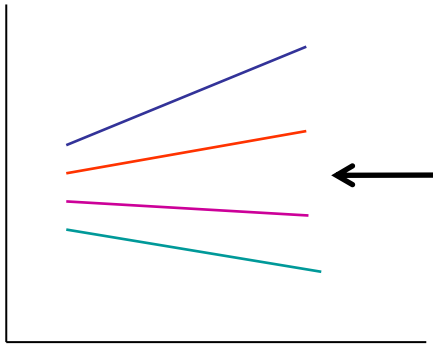
- **Between**-Person* (**BP**) Variation:
  - Macro – **Level-2** – "**INTER**-individual Differences" – Time-Invariant
  - All longitudinal studies that begin as cross-sectional studies have this
- **Within**-Person* (**WP**) Variation:
  - Micro – **Level-1** – "**INTRA**-individual Differences" – Time-Varying
  - Only longitudinal studies can provide this extra type of information!
- Longitudinal studies allow examination of **both types** of relationships simultaneously (and their interactions)
  - **Any** variable measured over time usually has both BP and WP variation
  - BP = more/less than other people; WP = more/less than usual
- **I will use **person**, but "between" units can be anything that is measured repeatedly (like animals, schools, countries…)*

A Longitudinal Data Continuum

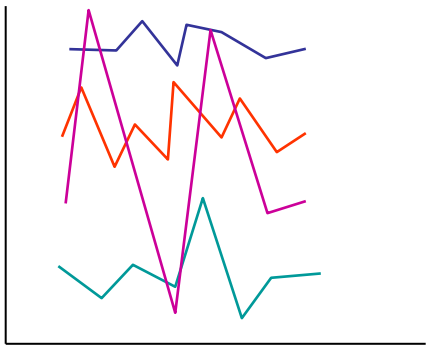
- **Within-Person (WP) Change**: Expect systematic effect(s) of time
  - e.g., "(Latent) Growth (Curve) Models" → **Time is meaningfully sampled**
  - If magnitude or direction of change differs across individuals, then the outcome's variance and covariance will change over time, too!

- **Within-Person (WP) Fluctuation**: Few *expected* effects of time
  - Outcome just varies/fluctuates over time (e.g., emotion, mood, stress)
  - **Time is just a way to get lots of data per person** (e.g., EMA studies)
  - Lends itself to questions about effects of relative changes and inconsistency

Pure WP Change



Pure WP Fluctuation



Sources of "Time" in Longitudinal Data

- What aspects of "**time**" are relevant in indexing change?
  - **WP change**: e.g., time in study, age, grade, time to/from event
  - **WP fluctuation**: e.g., time of day, day of week, day in study

- Does time vary **within persons (WP)** AND **between persons (BP)**?
  - If people differ in time at the study beginning (e.g., *accelerated* designs), we will need to **differentiate BP time effects from WP time effects**
  - If there is more than one kind of WP "time" (e.g., occasions within days), we will need to **differentiate distinct sources of WP time effects**

- Is time *balanced* or *unbalanced*?
  - **Balanced** = **shared** measurement schedule (not necessarily equal interval)
    - Although some people may miss some occasions, making their data "incomplete"
  - **Unbalanced** = people have **different** possible time values
    - By definition, the possible outcomes are at least partially "incomplete" across persons
    - This may be a consequence of using a time metric that also varies between persons

The Two Sides of *Any* Model

- **Model for the Means:**
  - *Aka* **Fixed Effects**, Structural Part of Model
  - What you are used to **caring about for testing hypotheses**
  - How the expected outcome for a given observation varies as a weighted function of its values of the predictor variables
    - Fixed slopes are **estimated constants** that multiply predictors

- **Model for the Variance (how many "piles"):**
  - *Aka* **Random Effects and Residuals**, Stochastic Part of Model
  - What you *were* used to **making assumptions about** instead
  - How **residuals are distributed and related** across sampling dimensions (persons, occasions) → these relationships are known as "**dependency**" and ***this is the primary way that longitudinal models differ from "regular" (GLM) regression models***

A Statistician's World View

- Outcome type: General (normal) vs. General*ized* (not normal)
- Dimensions of sampling: One (so one variance term per outcome) vs. **Multiple** (so multiple variance terms per outcome) → **OUR WORLD**

- **General Linear Models:** conditionally normal outcome distribution, **fixed effects** (identity link; only one dimension of sampling)

Note: OLS is only for GLM

- **General*ized* Linear Models:** **any conditional outcome distribution**, **fixed effects** through **link functions**, no random effects (one dimension)
- **General Linear Mixed Models:** conditionally normal outcome distribution, **fixed and random effects** (identity link, but multiple sampling dimensions)
- **General*ized* Linear Mixed Models:** **any conditional outcome distribution**, **fixed and random effects** through **link functions** (multiple dimensions)
  - Same concepts as for this course, but with more complexity in estimation

- "Linear" means fixed effects predict the *link-transformed* conditional mean of DV in a linear combination of (slope*predictor) + (slope*predictor)…

Multilevel Model (MLM) Word Salad

- MLM is the same as other terms you have heard of:
  - **Linear Mixed-Effects Model** (fixed + random effects, of which intercepts and slopes are specific kinds of effects)
  - **Random Coefficients Model** (because coefficients also = effects)
  - **Hierarchical Linear Model** (not same as hierarchical regression)

- Special cases of MLM:
  - Random Effects ANOVA or Repeated Measures ANOVA
  - (Latent) Growth Curve Model (where "Latent" implies SEM software)
    - Btw, most MLMs can be equivalently estimated as single-level SEMS
  - Within-Person Fluctuation Model (e.g., for EMA or daily diary data)
    - See also "dynamic" SEM or multilevel SEM (even without measurement models!)
  - Clustered/Nested Observations Model (e.g., for kids within schools)
    - If followed over time in same group, is "clustered longitudinal model"
  - Cross-Classified Models (e.g., teacher "value-added" models)
  - Psychometric Models (e.g., factor analysis, item response theory, SEM)

Options for Longitudinal Models

- Although models and software are logically separate, longitudinal data can be analyzed via multiple analytic frameworks (which are then usually tied to software):
  - "**Multilevel/Mixed-Effects/Hierarchical Linear Models**"
    - Dependency over time, persons, groups, etc. is modeled via random effects (multivariate through "levels" using stacked/long data)
    - Builds on GLM, generalizes easier to additional levels of analysis
  - "**Structural Equation Models**"
    - Dependency over time *only* is modeled via latent variables (single-level analysis using multivariate/wide data)
    - Generalizes easier to analysis of latent constructs, multivariate, mediation
  - Because random effects and latent variables are the same thing, many longitudinal models can be specified/estimated either way
    - And now "Multilevel Structural Equation Models" can do it all (maybe)…

What can MLM do for you?

1. **Model dependency across observations**
   - Longitudinal, clustered, and/or cross-classified data? No problem!
   - Tailor your model of sources of correlation to your data

2. **Include categorical or continuous predictors at any level**
   - Time-varying, person-level, group-level predictors for each variance
   - Explore reasons for dependency, don't just control for dependency

3. **Does not require same data structure for each person**
   - Unbalanced or missing data? No problem!

4. **You already know how (or you will soon)!**
   - Use SAS Mixed, STATA Mixed, R, SPSS Mixed, Mplus, HLM, MlwiN…
   - What's an intercept? What's a slope? What's a pile of variance?

1. Model Dependency

- Outcomes from the same sampling unit (i.e., person) will have one or more sources of **dependency → correlated residuals**

- **"Levels" for dependency** ="levels of random effects"
  - More generally, a "**level**" refers to a **dimension of sampling** that has **unexplained** outcome variability represented by 1+ random effects
  - Sampling dimensions can be **nested** in each other
    - e.g., time within person, time within group, trial within person
  - If you can't figure out the direction of your nesting structure, odds are good you have a **crossed sampling design** instead
    - e.g., persons crossed with items, raters crossed with targets
  - To have a "level", there must be random outcome variation due to sampling that **remains** after including the model's fixed effects
    - e.g., treatment vs. control does not create another level of "group" (but it would if you had multiple treatment and multiple control groups)

Longitudinal dependency comes from…

- **Mean differences** across sampling units (e.g., persons)
  - Creates constant dependency (residual correlation) over time
  - Will be represented by a **random intercept** in our models (or could be addressed by fixed main effects of person ID)
- **Slope differences** in effects of time-varying predictors
  - Individual differences in change over time, stress reactivity
  - Creates non-constant dependency, the size of which depends varies over time based on the value of the time-varying predictor
  - Will be represented by **random slopes** in our models (or could be addressed by fixed interaction effects of person ID)
- Non-constant within-person correlation for other unknown reasons (time-dependent autocorrelation)
  - Can add other patterns of correlation as needed for this (what I will call "alternative covariance structure" models)

Why should we care about dependency?

- In other words, what happens if we have the wrong model for the variance (i.e., assume independent residuals instead)?

- **Validity of the tests of the predictors** depends on having the "most right" (or least wrong) model for the variance
  - Estimates of slopes for level-specific predictors (i.e., only level 2 or level 1) will usually be ok (because they are from the model for the means)
  - Estimates of slopes for most time-varying predictors will NOT be ok!
  - **Standard errors** (and thus ***p*-values**) of slopes can be **incorrect**

- The sources of variation that are present in your outcome will dictate **what kinds of predictors** will be useful
  - Between-Person variation needs time-invariant predictors
  - Within-Person (time) variation needs time-varying predictors
  - Between-whatever variation needs Between-whatever predictors…
  - Bottom line: If it don't vary, GAME OVER for predicting it

2. Include categorical or continuous predictors at any level of analysis

- "ANOVA" tests differences among discrete categories measured once (Between) or repeatedly measured (RM)
- "Regression" tests slopes for quantitative predictors *measured once* per person on outcomes *measured once* per person
- What if a predictor is measured repeatedly but it can't be characterized by discrete "conditions"?
  - ANOVA or Regression won't work (easily) → you need MLM

- Some things don't change over time → time-invariant
- Some things do change over time → time-varying
- Some things are measured at higher levels
- Interactions are possible at same level or across levels

3. Does not require same data structure per person (by accident or by design)

RM ANOVA: uses **wide** data structure (one person per row):

ID	Sex	YT1	YT2	YT3	YT4
100	0	5	6	8	12
101	1	4	7	.	11

People missing any data are excluded (data from ID 101 are not included at all)

MLM: uses a **long** (stacked) data structure (one occasion per row):

ID	Sex	Time	Y
100	0	1	5
100	0	2	6
100	0	3	8
100	0	4	12
101	1	1	4
101	1	2	7
101	1	3	.
101	1	4	11

Only rows missing data are excluded

100 uses 4 cases

101 uses 3 cases

Time can also be **unbalanced** across people such that each person can have their own measurement schedule: Time = "0.9" "1.4" "3.5" "4.2"…

4. You already know how (or you will soon)!

- If you can do GLMs, you can do MLMs
  (*and if you can do generalized linear models,*
  you can do generalized multilevel models, too)

- Let's review: How do you interpret an estimate for…
  - the intercept?
  - the slope of a quantitative predictor?
  - the slope of a categorical predictor?
  - a variance component ("pile of variance")?